5 Lecture 5

5.1 Matrix multiplication

Lemma 5.1. *$T, S : V \to W$ be linear transformations. Let $\beta = (v_1, \ldots, v_n)$ be a basis of V and $\gamma = (w_1, \ldots, w_m)$ be a basis of W. Then $[T+S]_\beta^\gamma = [T]_\beta^\gamma + [S]_\beta^\gamma$ and $[cT]_\beta^\gamma = c[T]_\beta^\gamma$.*

Proof. Use linearity of $T, S, \phi_\beta, \phi_\gamma$ to verify. □

Corollary 5.2. *$\mathcal{L}(V, W) \cong M_{m\times n}(F)$ under the map $T \mapsto [T]_\beta^\gamma$. In particular,* $\dim \mathcal{L}(V, W) = (\dim V)(\dim W)$.

Proof. Show injectivity and surjectivity of the map $T \mapsto [T]_\beta^\gamma$ using β and γ are bases. □

Definition 5.3. Let $A \in M_{m\times n}(F)$ and $B \in M_{n\times p}(F)$. Then the **matrix multiplication** $C = AB \in M_{m\times p}(F)$ is given by

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}.$$

The i-th row, j-th column of C is given by multiplying the i-th row of A and j-th column of B.

Proposition 5.4 (Matrix multiplication is composition of linear transformation)**.** *Let $T : V \to W$ and $S : W \to U$ be linear transformations, and β, γ and δ are bases of V, W, U respectively. Then*

$$[ST]_\beta^\delta = [S]_\gamma^\delta \, [T]_\beta^\gamma.$$

Proof. We write $[S]_\gamma^\delta = (a_{ij}) \in M_{m\times n}(F)$ and $[T]_\beta^\gamma = (b_{ij}) \in M_{n\times p}(F)$ and $\beta = (v_1, \ldots, v_p)$, $\gamma = (w_1, \ldots, w_n)$ and $\delta = (u_1, \ldots, u_m)$. Then

$$\begin{aligned} ST(v_j) =& S\left(\sum_{k=1}^{n} b_{kj} w_k\right) \\ =& \sum_{k=1}^{n} b_{kj} S(w_k) \\ =& \sum_{k=1}^{n} b_{kj} \sum_{i=1}^{m} a_{ik} u_i \\ =& \sum_{i=1}^{m} \left(\sum_{k=1}^{n} a_{ik} b_{kj}\right) u_i. \end{aligned}$$

Thus $[ST]_\beta^\gamma = (c_{ij})$ where $c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$. □

We can translate definitions and results about linear transformations into matrices.

Definition 5.5. For $A \in M_{m\times n}(F)$, we define $L_A : F^n \to F^m$ to be the linear transformation $L_A(x) = Ax$. (We sometimes do not distinguish A and L_A.) The following notation are inherited from those of linear transformations
$\ker A = \ker L_A$,
$\operatorname{null} A = \operatorname{null} L_A$,
$\operatorname{im} A = \operatorname{im} L_A$,
$\operatorname{rank} A = \operatorname{rank} L_A$.

The following properties of matrix multiplication follows directly from Proposition 5.4 and Proposition 4.2 applied to L_A, L_B, L_C.

Proposition 5.6 (Matrix algebra rules)**.** *For any matrices of compatible sizes and any scalar $a \in F$:*
(i) $A(BC) = (AB)C$
(ii) $A(B + C) = AB + AC$
(iii) $(A + B)C = AC + BC$
(iv) $a(BC) = (aB)C = B(aC)$
(iv) $AI_n = A$, $I_m A = A$

The following rank nullity theorem for matrix follows from applying Corollary 4.14 to L_A.

Theorem 5.7. *Let $A \in M_{m\times n}(F)$. Then*

$$n = \operatorname{rank} A + \operatorname{null} A.$$

Definition 5.8. A matrix A is called **invertible** if L_A is invertible.

By Corollary 4.11, if A is invertible, then A is a square matrix. By Corollary 4.16, we have the following invertiblity criterion.

Corollary 5.9. *Let $A \in M_{n\times n}(F)$. The following are equivalent.*
(i) A is invertible.
(ii) $\operatorname{null} A = 0$.
(iii) $\operatorname{rank} A = n$.
(iv) There is $B \in M_{n\times n}(F)$ such that $AB = I$.
(v) There is $B \in M_{n\times n}(F)$ such that $BA = I$.
If any of the above holds, then the matrix B in (iv) (v) are the same and unique.

Definition 5.10. If A is invertible, then the B in (iv), (v) above are called the **inverse matrix** of A denoted by A^{-1}.

Corollary 5.11. *Let $T : V \to W$ be a linear transformation, β be a basis of V and γ be a basis of W. Then T is invertible if and only if $[T]^\gamma_\beta$ is invertible. Furthermore,*

$$[T^{-1}]^\beta_\gamma = ([T]^\gamma_\beta)^{-1}.$$

Proof. If T is invertible, then $I = [\mathrm{id}_V]^\beta_\beta = [T^{-1}T]^\beta_\beta = [T^{-1}]^\beta_\gamma [T]^\gamma_\beta$. Then $[T]^\gamma_\beta$ is invertible and we have $([T]^\gamma_\beta)^{-1} = [T^{-1}]^\beta_\gamma$.

If $A = [T]^\gamma_\beta$ is invertible, then since $\phi_\gamma T \phi_\beta^{-1} = L_A$, we have $T = \phi_\gamma^{-1} L_A \phi_\beta$. Since the right hand side are all invertible, we have T is invertible by Lemma 4.12. $\square$

5.2 Change of basis formula

Lemma 5.12. *Let $\beta = (v_1, \dots, v_n)$ and $\gamma = (w_1, \dots, w_n)$ be two bases of V. Let $P = (p_{ij}) = [\mathrm{id}]_\beta^\gamma \in M_{n\times n}(F)$. Then $v_j = \sum_{i=1}^m p_{ij} w_i$ for $1 \le j \le n$ and for any $v \in V$, $[v]_\gamma = P[v]_\beta$.*

Proof. This is a direct consequence of Theorem 4.22 with $T = \mathrm{id}$. □

Definition 5.13. The **change of coordinates matrix** from β-coordinates to γ-coordinates is the matrix $P = (p_{ij}) \in M_{n\times n}(F)$ such that $[v]_\gamma = P[v]_\beta$ for any $v \in V$. In other words, $P = [\mathrm{id}]_\beta^\gamma = ([v_1]_\gamma, \dots, [v_n]_\gamma)$.

The matrix P is also called the **change of basis matrix** (**transition matrix**) from γ to β in view of the property $v_j = \sum_{i=1}^n p_{ij} w_i$ for any $1 \le j \le n$. .

Remark 5.14. We note here that if P is the change of basis matrix from γ to β, then it is the change of coordinates matrix from β-coordinates to γ-coordinates. The role of β and γ are switched when we switch our view point from change of basis to change of coordinates.

Remark 5.15. We also note difference between the definition of change of basis matrix, $v_j = \sum_{i=1}^m p_{ij} w_i$ and the usual definition of a matrix A multiplying a vector x, $\sum_{j=1}^n a_{ij} x_j$. In the first case, the first index of the matrix P is summed over with w_j which is a vector, while in the second case, the second index of the matrix A is summed over with x_j which is a component of a vector. This is why some people will write $(v_1, \dots, v_n) = (w_1, \dots, w_n)P$ for the change of basis matrix thought of as the "row vector" $(w_1, \dots, w_n)$ multiplying the matrix P in the usual way.

Lemma 5.16. *Let $P = [\mathrm{id}]_\beta^\gamma$ be the change of basis matrix from γ to β. Then P is invertible and $P^{-1} = [\mathrm{id}]_\gamma^\beta$ is the change of basis matrix from β to γ.*

Proof. Since id is invertible and $\mathrm{id}^{-1} = \mathrm{id}$, by Corollary 5.11 we have $P = [\mathrm{id}]_\beta^\gamma$ is invertible and $P^{-1} = [\mathrm{id}]_\gamma^\beta$. □

Example 5.17. Let β be the standard basis of F^3 and $\gamma = ((1,0,0)^t, (1,1,0)^t, (1,1,1)^t)$, then the change of basis matrix from β to γ is $[\mathrm{id}]_\gamma^\beta = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$. In other words, $w_1 = v_1$, $w_2 = v_1 + v_2$, $w_3 = v_1 + v_2 + v_3$. To find the inverse matrix, we have $v_1 = w_1$, $v_2 = w_2 - v_1 = w_2 - w_1$, $v_3 = w_3 - v_1 - v_2 = w_3 - w_2$. Then $[\mathrm{id}]_\beta^\gamma = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$.

Theorem 5.18. *Let $T : V \to W$ be a linear transformation, β, β' be bases of V and γ, γ' be bases of W. Suppose $A = [T]_\beta^\gamma$, $B = [T]_{\beta'}^{\gamma'}$, $P = [\mathrm{id}_V]_{\beta'}^\beta$ and $Q = [\mathrm{id}_W]_{\gamma'}^\gamma$. Then we have*

$$B = Q^{-1}AP.$$

Proof. By Proposition 5.4, $B = [T]_{\beta'}^{\gamma'} = [\mathrm{id}]_{\gamma}^{\gamma'}[T]_{\beta}^{\gamma}[\mathrm{id}]_{\beta'}^{\beta} = Q^{-1}AP$. □

Corollary 5.19. *Let $T : V \to V$ be a linear transformation and β ,β' be two bases of V. Suppose $A = [T]_{\beta}^{\beta}$, $B = [T]_{\beta'}^{\beta'}$ and $P = [\mathrm{id}]_{\beta'}^{\beta}$ is the change of basis matrix from β to β'. Then*

$$B = P^{-1}AP.$$

Definition 5.20. Given $A, B \in M_{n\times n}(F)$, we say that A and B are **similar** if there is $P \in M_{n\times n}(F)$ invertible such that $B = P^{-1}AP$.

Finding the simplest matrix which is similar to a given $A \in M_{n\times n}(F)$ is the main reason to introduce eigenvalues and eigenvectors which will be discussed later in the notes. However, if one allows different bases of V and W as in Theorem 5.18, then we have the following theorem.

Theorem 5.21. *Let $T : V \to W$ be a linear transformation. Then there exists a basis β of V and a basis γ of W such that*

$$[T]_{\beta}^{\gamma} = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix},$$

where $r = \operatorname{rank} T$. In particular, the rank is the only invariant of a linear transformation under change of basis on both the domain and the target.

In matrix language, for any $A \in M_{m\times n}(F)$, then there exist invertible matrices $P \in M_{n\times n}(F)$ and $Q \in M_{m\times m}(F)$ such that $Q^{-1}AP = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$.

Proof. Let $v_1, \dots, v_n$ be a basis of V. Then $\operatorname{im} T = \operatorname{span}(T(v_1), \dots, T(v_n))$. By Lemma 2.11, $(T(v_1), \dots, T(v_n))$ contains a basis of $\operatorname{im} T$ which after relabeling the indices, we assume to be $T(v_1), \dots, T(v_r)$ for $r = \operatorname{rank} T$. In particular, $T(v_1), \dots, T(v_r)$ is linearly independent. By Corollary 2.16, we can extend it to be $\gamma = (T(v_1), \dots, T(v_r), w_{r+1}, \dots, w_m)$ which is a basis of W. Since $\operatorname{span}(T(v_1), \dots, T(v_r)) = \operatorname{im} T$, there exist $a_{ij} \in F$ for $1 \le i \le r$ and $r + 1 \le j \le n$ such that $T(v_j) = \sum_{i=1}^{r} a_{ij} T(v_i)$. We define $e_j = v_j - \sum_{i=1}^{r} a_{ij} v_i$ for $r+1 \le j \le n$. Apply Lemma 2.12 to $e_{r+1}, \dots, e_n$ one at a time, we see that $\beta = (v_1, \dots, v_r, e_{r+1}, \dots, e_n)$ is a basis of V. By definition of e_j, we have $T(e_j) = 0$ for $r+1 \le j \le n$. Then under the basis β and γ, we have $[T]_{\beta}^{\gamma} = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$. □

5.3 Dual space

Definition 5.22. A **linear function** (**linear functional/ 1-form/covector**) on V is a linear transformation $\varphi : V \to F$. The **dual space** V^* of V is the set of all linear functions on V. In other words $V^* = \mathcal{L}(V, F)$. In particular, V^* is a vector space. The **natural pairing** $\langle \cdot, \cdot \rangle : V^* \times V \to F$ is defined as $\langle \varphi, x \rangle := \varphi(x)$ for $x \in V, \varphi \in V^*$.

Remark 5.23. In quantum mechanics, a linear function is also called a bra-vector denoted by $\langle \varphi |$ while a usual vector is a ket-vector denoted by $| \psi \rangle$. To pair them, just write them together $\langle \varphi | \psi \rangle$ and it becomes a "bracket" which is the combined word of bra and ket.

Remark 5.24. When we don't have an inner product on a vector space, e.g. for vector spaces over finite fields, the natural paring serves as an "inner product". This is why we write it the same way as an inner product.

Example 5.25. For $V = F^n$, every linear functional $\varphi : F^n \to F$ is given by

$$\varphi \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = (a_1, \dots, a_n) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = a_1 x_1 + \cdots + a_n x_n.$$

for some row vector $(a_1, \dots, a_n) \in M_{1\times n}(F)$. Thus $(F^n)^*$ the space of row vectors while F^n is the space of column vectors.

Definition 5.26. Given a basis $\beta = (v_1, \dots, v_n)$ of V, the **dual basis** $\beta^* = (\varphi_1, \dots, \varphi_n)$ of V^* is defined by $\varphi_i(v_j) = \delta_{ij}$ for $1 \le i, j \le n$. By Lemma 4.21, φ_i is uniquely determined by its value at a basis.

Proposition 5.27. *The dual basis is a basis of V^*. In particular,* $\dim V^* = \dim V$.

Proof. Let $\varphi_1, \dots, \varphi_n$ be the dual basis of $\beta = (v_1, \dots, v_n)$. We first show that $\varphi_1, \dots, \varphi_n$ is linearly independent. Suppose $a_1\varphi_1 + \cdots + a_n\varphi_n = 0$. We evaluate both sides on v_i for $1 \le i \le n$ and get $a_1\varphi_1(v_i) + \cdots + a_n\varphi_n(v_i) = 0$. Since $\varphi_i(v_j) = \delta_{ij}$, we have $a_i = 0$ for all $1 \le i \le n$. Hence $\varphi_1, \dots, \varphi_n$ is linearly independent.

Next we show that $\mathrm{span}(\varphi_1, \dots, \varphi_n) = V^*$. Let $\varphi \in V^*$. For any $v \in V$ with $[v]_\beta = (x_1, \dots, x_n)^t$, we have $\varphi_i(v) = x_i$ for $1 \le i \le n$. We write $a_1 = \varphi(v_1), \dots, a_n = \varphi(v_n)$. Then by Lemma 4.21, $\varphi(v) = a_1x_1 + \cdots + a_nx_n = a_1\varphi_1(v) + \cdots + a_n\varphi_n(v) = (a_1\varphi_1 + \cdots + a_n\varphi_n)(v)$. Since this is true for any $v \in V$ we have $\varphi = a_1\varphi_1 + \cdots + a_n\varphi_n$. $\square$

Remark 5.28. From the proof above, we see that the coordinates of v under the basis β is given by $\phi_\beta(v) = (\varphi_1(v), \dots, \varphi_n(v))^t$.

5.4 Transpose of a linear transformation

Definition 5.29. Let $T : V \to W$ be a linear transformation. We define the **transpose map** (**dual map/pull-back**) $T^t : W^* \to V^*$ as $T^t(\varphi) = \varphi \circ T$. That is, for any $\varphi \in V^*$ and $v \in V$, $(T^t(\varphi))(v) = \varphi(T(v))$ or $\langle T^t\varphi, v\rangle = \langle \varphi, Tv\rangle$. Since $T^t(\varphi) = \varphi \circ T$, $T^t(\varphi)$ is a linear function on V since it is a composition of linear transformations. The map T^t is linear due to the algebraic rules of composition by linear transformations in Proposition 4.2.

Proposition 5.30. *Let $T : V \to W$ be a linear transformation. Let $\beta = (v_1, \dots, v_n)$ be a basis of V and $\gamma = (w_1, \dots, w_m)$ be a basis of W. Let β^* be the dual basis of β and γ^* be the dual basis of γ. Then $[T^t]_{\gamma^*}^{\beta^*} = ([T]_\beta^\gamma)^t$.*

Proof. Let $\gamma^* = (\psi_1, \ldots, \psi_m)$ be the dual basis of γ and $[T]_\beta^\gamma = (a_{ij})$. We want to find the coordinates of $T^t(\psi_i)$ under $\beta^* = (\varphi_1, \ldots, \varphi_n)$, the dual basis of β. Let $v \in V$ and $v = x_1 v_1 + \cdots + x_n v_n$. Note that $x_i = \varphi_i(v)$. Then we have $T^t(\psi_i)(v) = \psi_i(T(v)) = \psi_i(\sum_{k=1}^m (\sum_{j=1}^n a_{kj} x_j) w_k) = \sum_{j=1}^n a_{ij} x_j = \sum_{j=1}^n a_{ij} \varphi_j(v) = (\sum_{j=1}^n a_{ij} \varphi_j)(v)$. Thus $T^t(\psi_i) = \sum_{j=1}^n a_{ij} \varphi_j$. That is $[T^t(\psi_i)]_{\gamma^*} = \begin{pmatrix} a_{i1} \\ \vdots \\ a_{in} \end{pmatrix}$. Hence $[T^t]_{\gamma^*}^{\beta^*} = ([T]_\beta^\gamma)^t$. $\square$